Visualizing Industrial Dynamics via Stock Market

Visualizing Industrial Dynamics Through Represented Stock Capital Changes Over about the Past Decade.

Chun Yuan
u6052162@utah.edu, u6052162
Zixuan Lu
birdpeople1984@gmail.com, u6053162

Content

This page is intentionally left blank.

Project Proposal

1.1 Basic info

1. **Project title:** Visualize Dynamics of Stock Market
2. **Group members:**
   Chun Yuan, u6052162@utah.edu, u6052162;
   Zixuan Lu, birdpeople1984@gmail.com, u6053162.
3. **Repo:** https://github.com/dataviscourse2024/group-project-market-dynamic-visualization.

1.2 Background & Motivation

Analyzing the industrial structure offers an effective perspective for investment analysis. Over time spans ranging from several decades to a century, countless emerging industries have been born and have flourished, while older, less competitive industries have been phased out or downgraded. Observing and analyzing these trends provides valuable insights, allowing for predictions of economic cycles and offering macroeconomic investment guidance.

Industries are always evolving, driven by technological advances, consumer behavior, and regulations. New industries, like digital technology, renewable energy, and biotechnology, have created significant growth opportunities by challenging older business models and reshaping markets on the most recent decades. These sectors can offer high returns to investors who recognize and invest in them early. On the other hand, older industries that fail to adapt may decline, as seen with traditional manufacturing in many developed countries. Understanding which industries are likely to struggle can help investors adjust their portfolios to avoid potential losses.

By systematically monitoring these industrial shifts, investors can develop a more nuanced understanding of economic cycles. They can identify emerging opportunities and potential risks, and allocate their capital more efficiently. Such analysis is essential for formulating long-term investment strategies that align with evolving economic realities.

1.3 Project Objectives

In this project, we aim to provide an effective method for visualizing the dynamics of industrial structures. The major questions we wants to answer is:

1. What are the general patterns of the market (economic cycle patterns)?
2. Do industries developments reflect economic cycle patterns (cyclical/counter-cyclical industry)?
3. Are industreis growth balanced over a limited period?

4. Are there significant events that have had observable impacts on certain industries?

We believe these four questions are crucial for understanding how industries perform relative to the overall economic cycle at a limit time-span and for making informed investment decisions.

Considerations for Project Execution and Progress:

Note. Rather than creating our own classification standards, we will use the **Global Industry Classification Standard (GICS)**, which is widely adopted in the industry[1, 2]. There are numerous economic indicators that can reflect changes in industries, but we want to keep our approach straightforward and focused. Given that **stock market capitalization** is a direct indicator of industrial trends, preferences, and the development of companies and industries, we will use the **market capitalizations of a few representative stocks within each sector** as our key metric. This will allow us to visualize how their market values have changed over a certain period.

Since this is a short-term project designed for a single semester, we will limit the time frame of the data we visualize to **the most recent 5-10 years**, essentially covering up to a decade. This time frame will provide us with a manageable yet meaningful period for observing significant trends and shifts within various industries.

By focusing on a concise time span and using a clear, widely recognized standard like GICS, we believe our visualization will offer a practical tool for understanding industry dynamics without overwhelming complexity.

1.4 Data

1.4.1 GICS Industrial Sector and Industry

The Global Industry Classification Standard (GICS) is a widely recognized system for categorizing companies into specific sectors and industries, developed by MSCI and S&P Global in 1999, providing a consistent framework for analyzing companies worldwide, allowing investors and analysts to compare performance and trends across various sectors and industries. It divides the global market into 11 sectors, 24 industry groups, 69 industries, and 158 sub-industries, covering the full spectrum of economic activities.

For this project, we will use GICS to classify companies and industries with sector and industry tag. To be specific, we will select the stock based on this classification. The stock will be clustered respected to such tag. A group of stock within the same industrial tag will be grouped up to reflect the development of such industry. We will use 2-level of the classification information, which means specify to GICS industry group (omitting industies-level and sub-industries-level). Such classification information can be found at GICS white paper book.

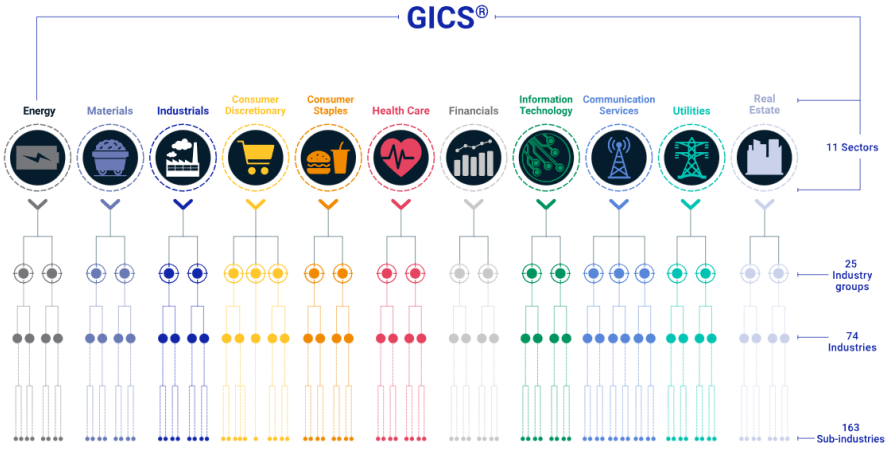


Figure 1.1: GICS Overview.

1.4.2 Stock Code with Tag

On the next stage, we will obtain the stock code list based on GICS Industrial sector and industry, which means we will select a group of represented stock with stock code based on GICS classification. Fortunately, we found a public data based on S&P 500 (2022), which collects totally 512 stock with {Ticker Exchange, GICS Sector, GICS Industry Group, GICS Industry, GICS Sub-Industry} as meta-data. This data can be found at Here.

Ticker	Exchange	GICS Sector	GICS Industry Group	GICS Industry	GICS Sub-Industry

Figure 1.2: Attribute: stock code based on GICS.

1.4.3 Stock meta-data

Finally, with the tagged stock code list, we will query a public database to obtain the meta-data of each single stock. The meta-data we queried for each stock will contain closing pric, market capitalization, trading volume, etc, combined with its stock code and GICS tag.

The public open source stock meta-data database we are going to query is: AKShare. We will use 5-10 years (end for 2024.8) with week-level interval data. This meta-data has to be query within Python. On the next section we will show the data processing strategy to format our data.

1.5 Data Processing

We will collect and manipulate our data to form a final format. Each entry of the data will have these attributes:

1. Time (quantity): date of the stock
2. Stock code (category): the code of the represented stock
3. Closing price (quantity): the closing price of the stock
4. Market capitalization (quantity): the market capitalization of the stock
5. Trading volume (quantity): the trade volume of the stock

6. GICS sector tag (category): one of the 11 GICS sector tag

7. GICS industry tag (category): one of the 25 GICS industry tag

1.6 Visualization Design

1.6.1 Overview

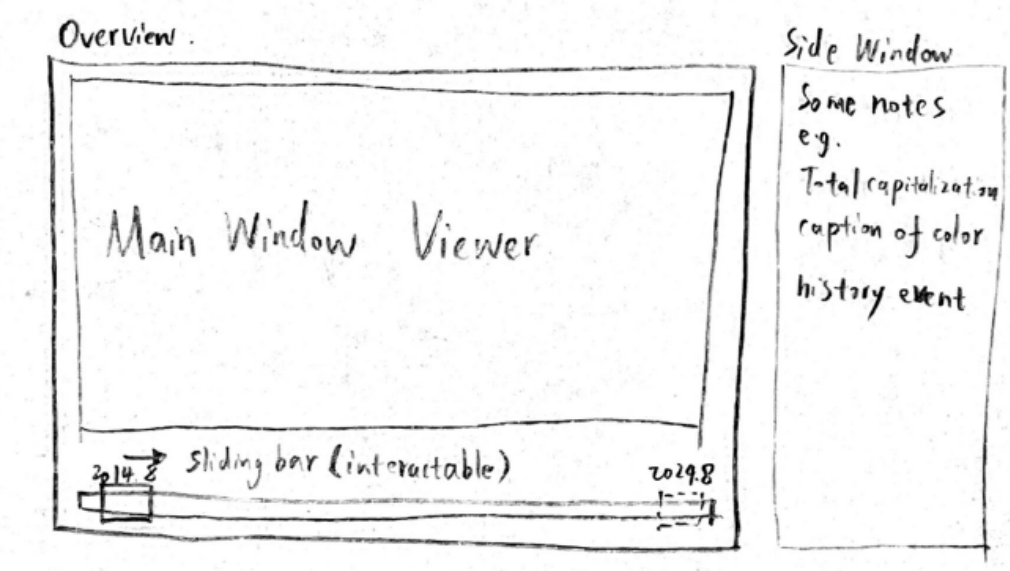


Figure 1.3: Overview of the visualization.

Figure 1.3 provides an overview of our visualization interface. The main interface consists of a side panel for displaying notes, an interactive sliding bar (timeline), and a central viewer for visualization. As the user adjusts the timeline slider, the content in the main viewer dynamically updates to reflect the selected time frame. The side panel presents relevant text, such as important historical events and other contextual information.

1.6.2 Specific design on main windows

The main window will display stock prices corresponding to each GICS sector or industry to illustrate the evolution of a given industry. It will present both the performance of individual stocks and the total market capitalization of the entire industry. To achieve this, we have developed three potential solutions:

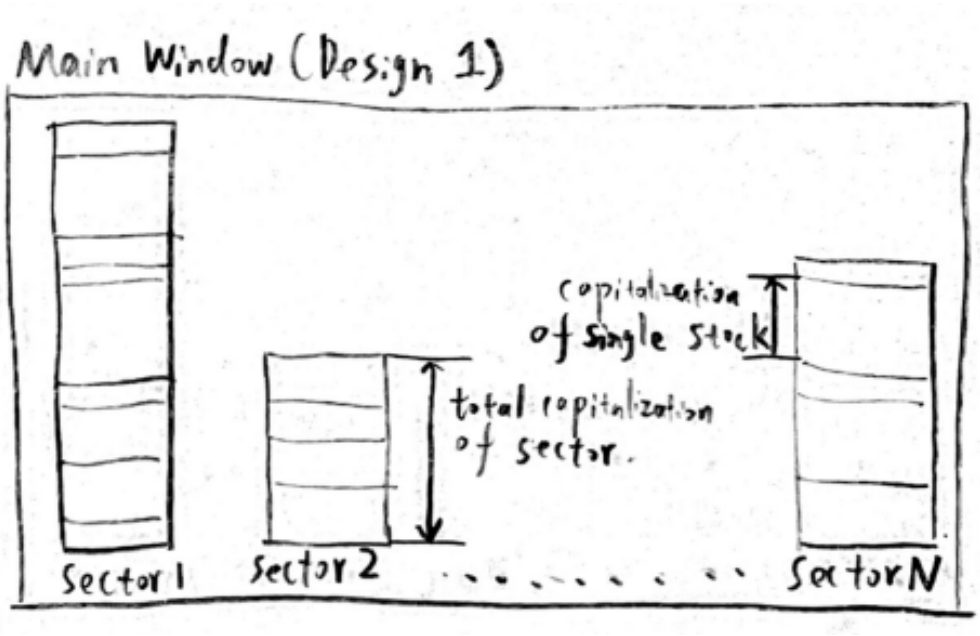


Figure 1.4: Design 1: Using stacked bar diagram.

The first idea we considered was using a traditional stacked bar chart(Fig. 1.4) to visualize the data. However, this approach was discarded for two main reasons. First, if a single industry grows rapidly, the axis may change significantly, leading to an axis-compression problem, where a large maximum scale can obscure smaller values. Second, each industry or sector may include stocks of varying sizes, which can create issues when stacking them on a unified scale. Despite these drawbacks, this design does effectively display both the total market capitalization and the values of individual stocks.

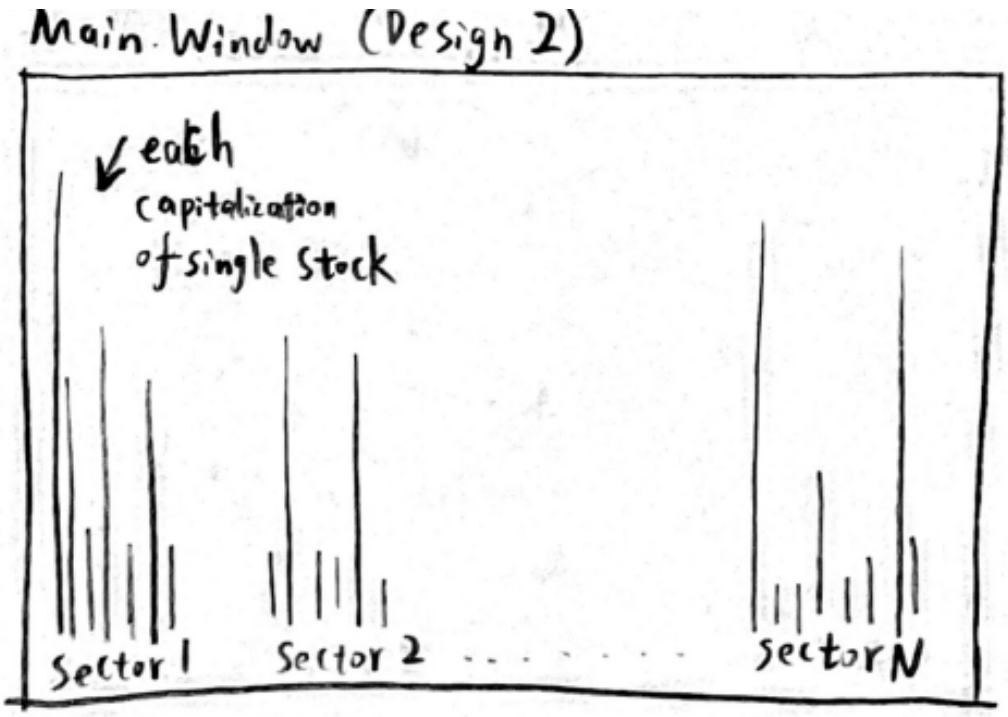


Figure 1.5: Design 2: Using separated bar diagram.

Another idea is to visualize the data using separate bar charts arranged horizontally (Fig. 1.5), with each chart representing the price of an individual stock. While this design effectively highlights the specific value of each stock, it struggles to convey the total market capitalization of each sector or industry. The total capitalization is represented by the combined area of all the horizontal bars. However, this design gave us the insight that using area as a key factor can effectively illustrate relative quantities in a visually intuitive and compelling way.

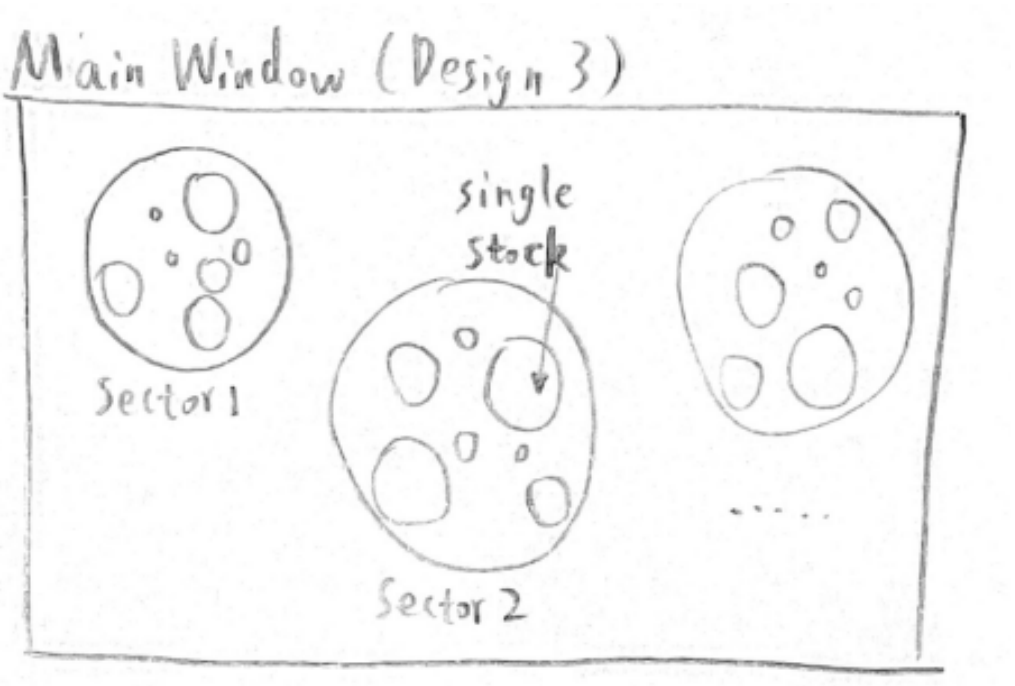


Figure 1.6: Design final: Using bubble diagram.

The final design involves using a clustered circle visualization to represent each sector's or industry's total capitalization, with smaller circles within each cluster representing individual stock prices (see Fig. 1.6). We believe this design offers two main advantages: First, it efficiently displays the relative value of each sector, as sectors or industries are naturally grouped together. Second, it clearly distinguishes each individual stock within the cluster. Additionally, by using circles to represent data, we eliminate the issue associated with using a vertical axis with a unified scale. Smaller stocks appear as points on the display, while more developed stocks are represented by larger circles. This dynamic representation is also intuitive as the user adjusts the timeline(see Fig. 1.7).

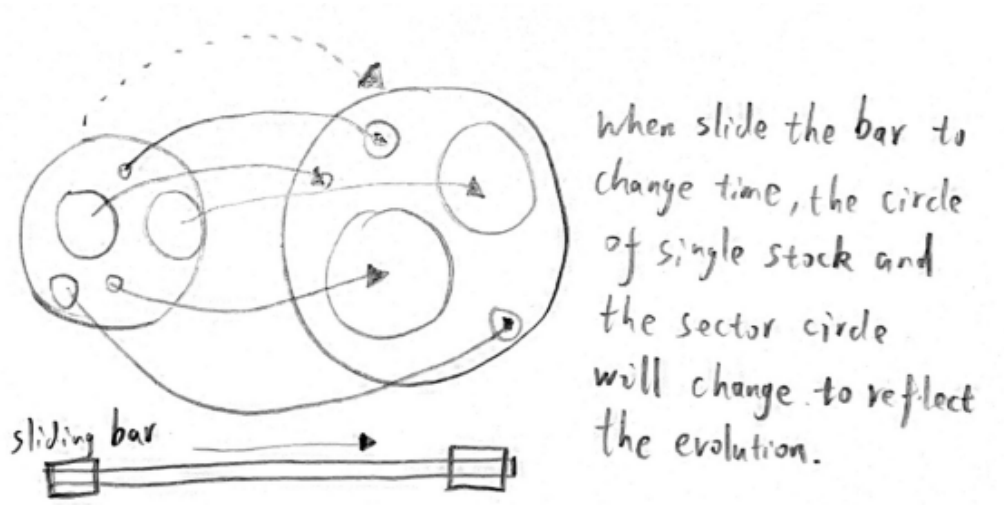


Figure 1.7: Dynamics of bubble diagram when sliding the time bar.

We also provided an stretch up of our optional feature (see Fig. 1.8): as the time bar is adjusted, a separate window will display the total capitalization value of each sector or industry using a line chart.

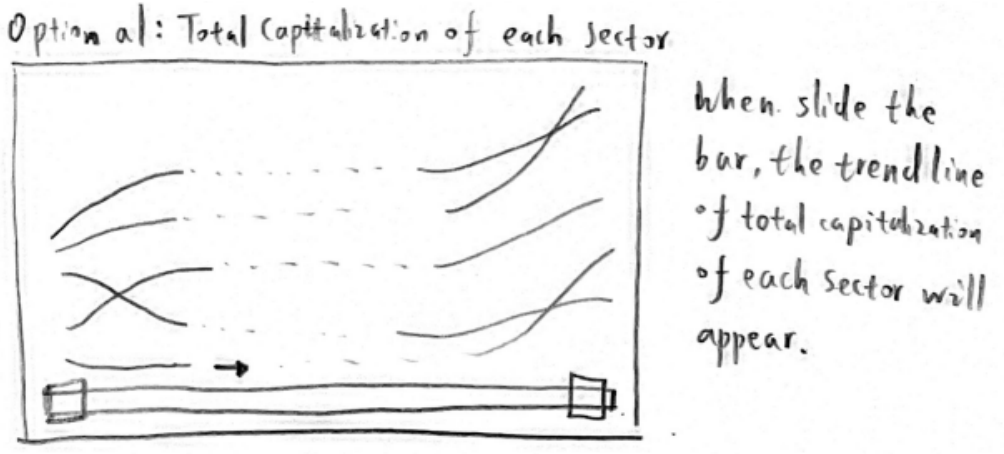


Figure 1.8: Optional feature: display the total capitalization.

1.7 Must-Have Features

Here we list some of the must have features:

1. On the bottom of the visualization windows, there should be a sliding timeline, which allow users to interact with the windows to show different time period of the market.

2. The main window can display clustered data, each cluster are based on GICS sector/industry classifications.

3. On each cluster, there should be a set of circles to represent single stock. The diameter of the circles should reflect the price of such stock.

1.8 Option Features

Here we list some optional features.

1. Another separate window displays the total market capitalization of each sector and industry (summation of single stock's capitalization within it), reflecting the overall economic development situation.

2. Another separate window displays the percentage of each sector or industry relative to the total market capitalization (summation of all).

1.9 Project Schedule

Our schedule:

1. Week 5-6: Finish the data collection and data processing.

2. Week 7-8: Add a sliding timeline, which allow the user to interact with. Main window can reflect the action of such time.

3. Week 9: Adding circles to display the stock data.

4. Week 10: Finalize the project to make it beautiful.

Data Processing

This chapter is filled on Oct.19, 2024.

2.1 The meta-data required

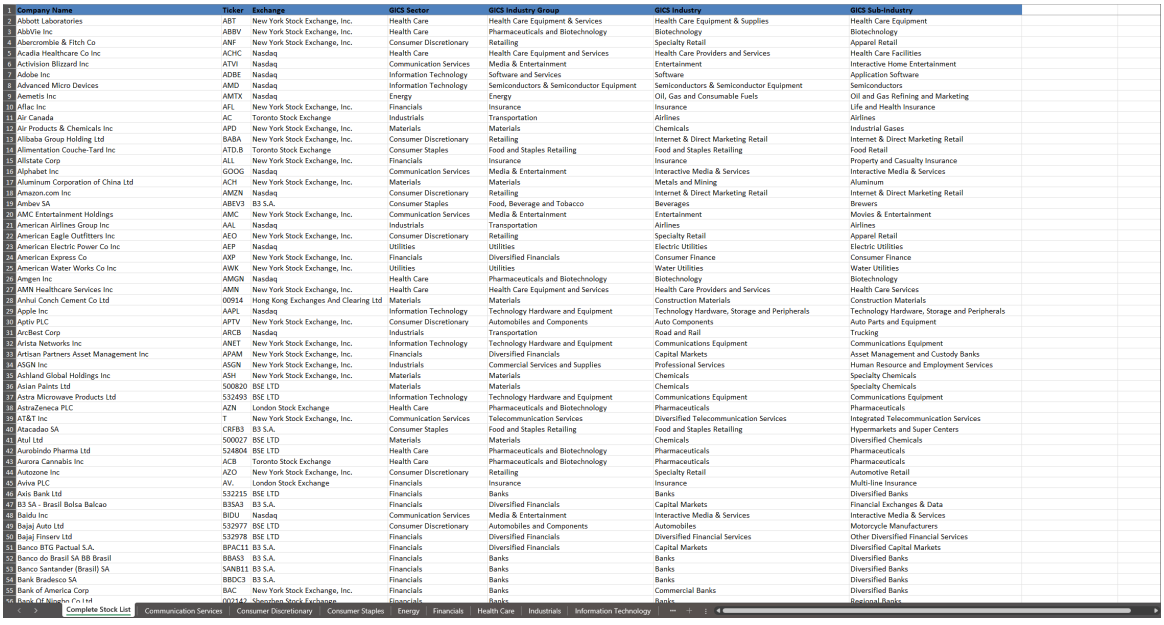
In the previous section of the project proposal, we confirmed that this project are going to use the U.S. stock market capitalization to reflect industry development trends and capital flow (using GICS industry classification standard). Therefore, we first need to collect and manipulate the data.

Specifically, the following data is required:

1. GICS industry classification stock codes
2. The meta-data of all stocks through these codes
3. Stock prices calculated from the meta-data
4. (Optional) Other clustering data, such as the total value of a specific industry

2.2 The Crawler and feeding data structure

Firstly, we researched GICS industry classification data and used data from here to obtain the stock codes classified by GICS sector and industry categories within the S&P 500 index. At the same time, we made data cleaning to narrow to the U.S. market stocks.



Company Name	Ticker	Exchange	GICS Sector	GICS Industry Group	GICS Industry	GICS Sub-Industry
Abbott Laboratories	ABT	New York Stock Exchange, Inc.	Health Care	Health Care Equipment & Services	Health Care Equipment & Supplies	Health Care Equipment
AbbVie Inc	ABBV	New York Stock Exchange, Inc.	Health Care	Pharmaceuticals and Biotechnology	Biotechnology	Biotechnology
Abercrombie & Fitch Co	ANF	New York Stock Exchange, Inc.	Consumer Discretionary	Retailing	Specialty Retail	Apparel Retail
Acadia Healthcare Co Inc	ACHC	Nasdaq	Health Care	Health Care Equipment and Services	Health Care Providers and Services	Health Care Facilities
Activision Blizzard Inc	ATVI	Nasdaq	Communication Services	Media & Entertainment	Entertainment	Interactive Home Entertainment
Adobe Inc	ADBE	Nasdaq	Information Technology	Software and Services	Software	Application Software
Advanced Micro Devices	AMD	Nasdaq	Information Technology	Semiconductors & Semiconductor Equipment	Semiconductors & Semiconductor Equipment	Semiconductors
Aemetis Inc	AMTX	Nasdaq	Energy	Energy	Oil, Gas and Consumable Fuels	Oil and Gas Refining and Marketing
Aflac Inc	AFL	New York Stock Exchange, Inc.	Financials	Insurance	Insurance	Life and Health Insurance
Air Canada	AC	Toronto Stock Exchange	Industrials	Transportation	Airlines	Airlines
Air Products & Chemicals Inc	APD	New York Stock Exchange, Inc.	Materials	Materials	Chemicals	Industrial Gases
Alibaba Group Holding Ltd	BABA	New York Stock Exchange, Inc.	Consumer Discretionary	Retailing	Internet & Direct Marketing Retail	Internet & Direct Marketing Retail
Alimentation Couche-Tard Inc	ATD.B	Toronto Stock Exchange	Consumer Staples	Food and Staples Retailing	Food and Staples Retailing	Food Retail
Allstate Corp	ALL	New York Stock Exchange, Inc.	Financials	Insurance	Insurance	Property and Casualty Insurance
Alphabet Inc	GOOG	Nasdaq	Communication Services	Media & Entertainment	Interactive Media & Services	Interactive Media & Services
Aluminum Corporation of China Ltd	ACH	New York Stock Exchange, Inc.	Materials	Materials	Metals and Mining	Aluminum
Amazon.com Inc	AMZN	Nasdaq	Consumer Discretionary	Retailing	Internet & Direct Marketing Retail	Internet & Direct Marketing Retail
Ambev SA	ABEV3	B3 S.A.	Consumer Staples	Food, Beverage and Tobacco	Beverages	Brewers
AMC Entertainment Holdings	AMC	New York Stock Exchange, Inc.	Communication Services	Media & Entertainment	Entertainment	Movies & Entertainment
American Airlines Group Inc	AAL	Nasdaq	Industrials	Transportation	Airlines	Airlines
American Eagle Outfitters Inc	AEO	New York Stock Exchange, Inc.	Consumer Discretionary	Retailing	Specialty Retail	Apparel Retail
American Electric Power Co Inc	AEP	Nasdaq	Utilities	Utilities	Electric Utilities	Electric Utilities
American Express Co	AXP	New York Stock Exchange, Inc.	Financials	Diversified Financials	Consumer Finance	Consumer Finance
American Water Works Co Inc	AWK	New York Stock Exchange, Inc.	Utilities	Utilities	Water Utilities	Water Utilities
Amgen Inc	AMGN	Nasdaq	Health Care	Pharmaceuticals and Biotechnology	Biotechnology	Biotechnology
AMN Healthcare Services Inc	AMN	New York Stock Exchange, Inc.	Health Care	Health Care Equipment and Services	Health Care Providers and Services	Health Care Services
Anhui Conch Cement Co Ltd	00914	Hong Kong Exchanges And Clearing Ltd	Materials	Materials	Construction Materials	Construction Materials
Apple Inc	AAPL	Nasdaq	Information Technology	Technology Hardware and Equipment	Technology Hardware, Storage and Peripherals	Technology Hardware, Storage and Peripherals
Aptiv PLC	APTV	New York Stock Exchange, Inc.	Consumer Discretionary	Automobiles and Components	Auto Components	Auto Parts and Equipment
ArcBest Corp	ARCB	Nasdaq	Industrials	Transportation	Road and Rail	Trucking
Arista Networks Inc	ANET	New York Stock Exchange, Inc.	Information Technology	Technology Hardware and Equipment	Communications Equipment	Communications Equipment
Artisan Partners Asset Management Inc	APAM	New York Stock Exchange, Inc.	Financials	Diversified Financials	Capital Markets	Asset Management and Custody Banks
ASGN Inc	ASGN	New York Stock Exchange, Inc.	Industrials	Commercial Services and Supplies	Professional Services	Human Resource and Employment Services
Ashland Global Holdings Inc	ASH	New York Stock Exchange, Inc.	Materials	Materials	Chemicals	Specialty Chemicals
Asian Paints Ltd	500820	BSE LTD	Materials	Materials	Chemicals	Specialty Chemicals
Astra Microwave Products Ltd	532493	BSE LTD	Information Technology	Technology Hardware and Equipment	Communications Equipment	Communications Equipment
AstraZeneca PLC	AZN	London Stock Exchange	Health Care	Pharmaceuticals and Biotechnology	Pharmaceuticals	Pharmaceuticals
AT&T Inc	T	New York Stock Exchange, Inc.	Communication Services	Telecommunication Services	Diversified Telecommunication Services	Integrated Telecommunication Services
Atacadao SA	CRFB3	B3 S.A.	Consumer Staples	Food and Staples Retailing	Food and Staples Retailing	Hypermarkets and Super Centers
Atul Ltd	500027	BSE LTD	Materials	Materials	Chemicals	Diversified Chemicals
Aurobindo Pharma Ltd	524804	BSE LTD	Health Care	Pharmaceuticals and Biotechnology	Pharmaceuticals	Pharmaceuticals
Aurora Cannabis Inc	ACB	Toronto Stock Exchange	Health Care	Pharmaceuticals and Biotechnology	Pharmaceuticals	Pharmaceuticals
Autozone Inc	AZO	New York Stock Exchange, Inc.	Consumer Discretionary	Retailing	Specialty Retail	Automotive Retail
Aviva PLC	AV.	London Stock Exchange	Financials	Insurance	Insurance	Multi-line Insurance
Axis Bank Ltd	532215	BSE LTD	Financials	Banks	Banks	Diversified Banks
B3 SA - Brasil Bolsa Balcao	B3SA3	B3 S.A.	Financials	Diversified Financials	Capital Markets	Financial Exchanges & Data
Baidu Inc	BIDU	Nasdaq	Communication Services	Media & Entertainment	Interactive Media & Services	Interactive Media & Services
Bajaj Auto Ltd	532977	BSE LTD	Consumer Discretionary	Automobiles and Components	Automobiles	Motorcycle Manufacturers
Bajaj Finserv Ltd	532978	BSE LTD	Financials	Diversified Financials	Diversified Financial Services	Other Diversified Financial Services
Banco BTG Pactual S.A.	BPAC11	B3 S.A.	Financials	Diversified Financials	Capital Markets	Diversified Capital Markets
Banco do Brasil SA BB Brasil	BBAS3	B3 S.A.	Financials	Banks	Banks	Diversified Banks
Banco Santander (Brasil) SA	SANB11	B3 S.A.	Financials	Banks	Banks	Diversified Banks
Bank Bradesco SA	BBDC3	B3 S.A.	Financials	Banks	Banks	Diversified Banks
Bank of America Corp	BAC	New York Stock Exchange, Inc.	Financials	Banks	Commercial Banks	Diversified Banks
Bank Of Ningbo Co Ltd	002142	Shenzhen Stock Exchange	Financials	Banks	Banks	Regional Banks

Complete Stock List | Communication Services | Consumer Discretionary | Consumer Staples | Energy | Financials | Health Care | Industrials | Information Technology

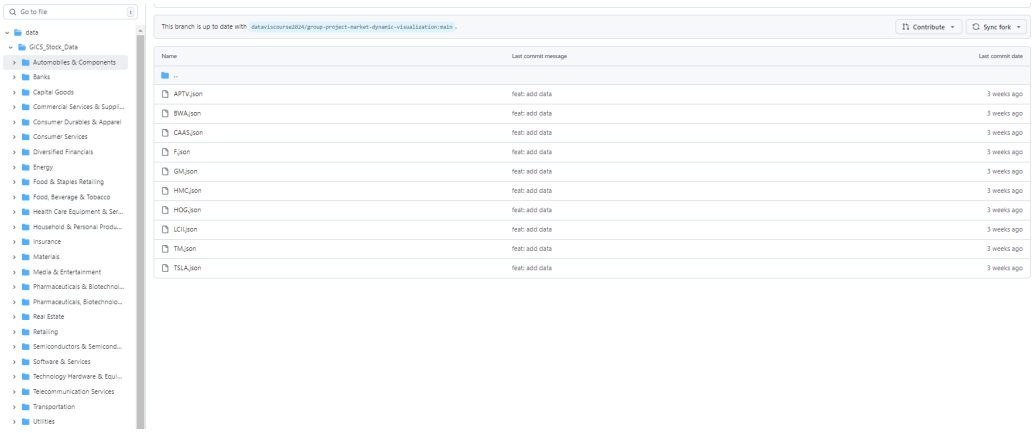


Figure 2.1: The data storage structure.


```
[
    {
        "date":"2017-12-08",
        "open":102.39,
        "close":83.24,
        "high":102.39,
        "low":81.89,
        "volume":11854567,
        "turnover":1023377226.0,
        "amplitude":136.48,
        "price_change_percent":454.19,
        "price_change_amount":68.22,
        "turnover_rate":4.46
    },
    {
        "date":"2017-12-15",
        "open":84.32,
        "close":82.42,
        "high":84.96,
        "low":80.99,
        "volume":13448660,
        "turnover":1142879988.0,
        "amplitude":4.77,
        "price_change_percent":-0.99,
        "price_change_amount":-0.82,
        "turnover_rate":5.06
    },
    {
        "date":"2017-12-22",
        "open":82.6,
        "close":82.81,
        "high":84.93,
        "low":82.15,
        "volume":8125324,
        "turnover":694842682.0,
        "amplitude":3.37,
        "price_change_percent":0.47,
        "price_change_amount":0.39,
        "turnover_rate":3.06
```


Figure 2.2: The meta-data for each stock.

Further, we used the AKShare API to collect the interval data of these stocks from 2017 to 2024. We stored these meta-data in a JSON file based on GICS industry classification and processed them with database scripts (e.g., JavaScript) for our webpage usage.

This page is intentionally left blank.

Bibliography

[1] MSCI Inc. and S&P Global. Global Industry Classification Standard (GICS®), 1999.

[2] State Street Global Advisors. GICS Sector and Industry Map, 2024.

[3] MSCI Inc. Global Industry Classification Standard (GICS®), 2024.